



The National DNA Data Bank of Canada: a Quebecer perspective

Emmanuel Milot¹, Marie M. J. Lecomte^{2,3}, Hugo Germain⁴ and Frank Crispino^{4*}

¹ Groupe de Recherche PRIMUS, Faculté de Médecine, Université de Sherbrooke, Sherbrooke, QC, Canada

² Mt. Albert Science Centre, The Institute of Environmental Science and Research Ltd., Auckland, New Zealand

³ The Department of Forensic Science, School of Chemical Sciences, The University of Auckland, Auckland, New Zealand

⁴ Département de Chimie, Biochimie et Physique, Université du Québec à Trois-Rivières, QC, Canada

Edited by:

Franco Taroni, University of Lausanne, Switzerland

Reviewed by:

Kesheng Wang, East Tennessee

State University, USA

Hsin-Chou Yang, Academia Sinica,

Taiwan

*Correspondence:

Frank Crispino, Département de Chimie, Biochimie et Physique, Université du Québec à Trois-Rivières, 3351 Boulevard des Forges, CP 500, Trois-Rivières, QC G9A 5H7, Canada
e-mail: frank.crispino@uqtr.ca

The Canadian National DNA Database was created in 1998 and first used in the mid-2000. Under management by the RCMP, the National DNA Data Bank of Canada offers each year satisfactory reported statistics for its use and efficiency. Built on two indexes (convicted offenders and crime scene indexes), the database not only provides increasing matches to offenders or linked traces to the various police forces of the nation, but offers a memory repository for cold cases. Despite these achievements, the data bank is now facing new challenges that will inevitably defy the way the database is currently used. These arise from the increasing power of detection of DNA traces, the diversity of demands from police investigators and the growth of the bank itself. Examples of new requirements from the database now include familial searches, low-copy-number analyses and the correct interpretation of mixed samples. This paper aims to develop on the original way set in Québec to address some of these challenges. Nevertheless, analytic and technological advances will inevitably lead to the introduction of new technologies in forensic laboratories, such as single cell sequencing, phenotyping, and proteomics. Furthermore, it will not only request a new holistic/global approach of the forensic molecular biology sciences (through academia and a more investigative role in the laboratory), but also new legal developments. Far from being exhaustive, this paper highlights some of the current use of the database, its potential for the future, and opportunity to expand as a result of recent technological developments in molecular biology, including, but not limited to DNA identification.

Keywords: DNA database, Canada, Québec, genetic engineering, forensic challenges

THE CANADIAN NATIONAL DNA DATA BANK

At the time the UK launched its DNA database in 1995, the exonerations of two wrongly accused individuals (Morin case, 1985 and Milgaard case, 1969) and the implementation of the C-104 bill (to amend the Criminal Code and the Young Offenders Act) acknowledged the need for a similar requirement in Canada and initiated the creation of the Canadian National DNA database (NDDB) by the Identification Act (Law C-37 of Dec. 10th, 1998) (Curran, 1997).

Following a nation-wide consultation with various institutional bodies (such as the Privacy Commissioner, the Canadian Bar Association, and the Canadian Police association), to address ethical, legal, and social implications issues also tackled by the National Human Genome Research Institute's (NHGRI) during the human Genome Project, its operative use was launched immediately after the proclamation of the S-10 bill on June 2000. A number of amendments led the NDDB to store genetic traces collected at crime scenes in the Crime Scene Index (CSI) and, under court order, the DNA profiles of offenders serving any sentence of imprisonment, for various categories of offences designated in section 487.04 of the criminal code, in the Convicted Offenders Index (COI).

Under the supervision of the DNA Data Bank Advisory Committee, composed of seven authoritative personalities involved in forensic biology, human rights and laboratory management, the NDDB is operated by the Royal Canadian Mounted Police (RCMP) for the benefit of all law enforcement agencies in the country, be it federal (the RCMP), provincial [the Ontario Police force or Sûreté du Québec (SQ)] or urban (depending on the level of police a town has to deliver in regard to its population), as provided by the RCMP at provincial and urban levels if requested.

The CSI is maintained by the RCMP labs, the Center of Forensic Sciences in Toronto (CFS¹) and the Laboratoire de sciences judiciaires et de médecine légale du Québec in Montréal (LSJML²) (Figure 1). On July 15th, 2013, the COI contains more than 273,000 profiles, while the CSI is nearing 87,000³.

Based on 13 DNA markers, DNA profiles are managed and compared using the Combined DNA Index System (CODIS). On a yearly basis, a RCMP report on the management of the NDDB

¹http://www.mcscs.jus.gov.on.ca/english/centre_forensic/CFS_intro.html.

²<http://www.securitepublique.gouv.qc.ca/lajml.html>.

³<http://www.rcmp-grc.gc.ca/nddb-bndg/stats-eng.htm>.

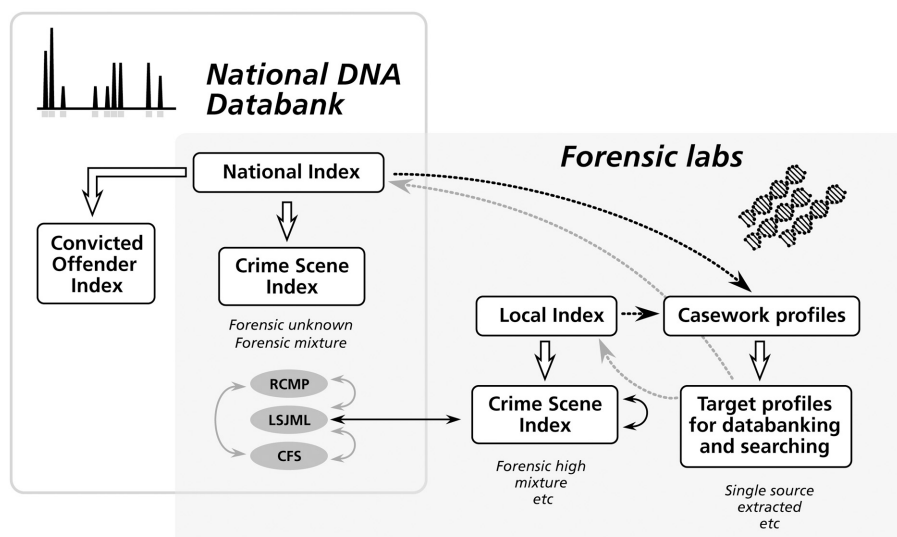


FIGURE 1 | The architecture of the National DNA Data Bank in Canada and its relationship to forensic laboratories. The bank is a national repository composed of two indexes. The Convicted Offender Index is managed centrally at the national level, while the National Crime Scene Index (CSI-nat) is managed collectively by three forensic labs (RCMP, LSJML, and CFS), with each lab being responsible for the profiles generated under its jurisdiction. The CSI-nat allows for inter-jurisdictional comparisons of crime scene profiles (solid gray arrows). The Local Crime Scene Index (CSI-loc)

corresponds to the databases maintained locally by forensic labs and containing DNA profiles that do not meet the criteria to be deposited in the NDDB (e.g., some complex mixtures). Local comparisons can be made both between profiles stored in the same CSI-loc and between profiles of the CSI-loc of a given lab and the portion of the CSI-nat managed by the same lab (solid black arrows; see **Figure 2**). Gray dotted arrows show the deposition of DNA profiles from caseworks into the national and local indexes, while black dotted arrows illustrate match information returned to the forensic labs.

provides statistics, financial costs information, a user guide for the reader, as well as information on the changing legal frame of the database, under the auspices of the Advisory Committee (Police, 2012).

The Advisory Committee controls and actively searches and suggests legislative and regulatory changes. This transparency in the management of the NDDB is what leads to the efficiency of the Canadian system, qualified as having “an astonishing degree of consistency in sampling regimes throughout the history of the Canadian DNA database.” (Walsh, 2009). Using the ratio of hits over the product of NC, (N being the numbers of profiles in the COI and C the numbers of profiles in the CSI), to assess the efficiency of DNA databases between four western countries (USA, the United Kingdom, the Netherlands, and New Zealand), the performance of the NDDB ranks just below New Zealand and is quite good, accounting for the lower proportion of the population being present in the database (0.5% for Canada instead of 2.1% for New Zealand). In regards to the public perception of civil rights, it could easily be deemed highly efficient. At least “Canada had a well-understood and effectively resourced concept of operation in place prior to the initiation of databasing” (Walsh, 2009). On such ground, Canada seems better prepared than many other countries to tackle new challenges facing forensic DNA identification.

THE LSJML DATABANKING STRATEGY

As with other DNA databanks, the NDDB holds key figures to address interpretation issues (Foreman et al., 2003; Dror and Hampikian, 2011) such as low copy numbers (LCN) (Lowe et al.,

2002; Phipps and Petricevic, 2007), mixed samples (Bill et al., 2005; Curran, 2008), and familial searches (Bieber et al., 2006; Reid et al., 2008; Miller, 2010; Murphy, 2010; Gershaw et al., 2011; Meyers et al., 2011; Pham-Hoi et al., 2013). While addressing the issue of familial searches is not yet on the agenda, as it would require changes to the Canadian legislation, LCN has become a routine challenge faced by forensic labs nationwide. Indeed, due to technological improvements, the detection of ever-smaller traces of DNA is now possible (Kayser and de Knijff, 2011). However, because of stochastic effects (drop-outs, drop-ins), this comes at the cost of lower repeatability and overall completeness of genetic profiles recovered from small quantities of DNA. This problem is made worse with mixtures owing to competitive amplification. Deconvoluting the information and sorting out the alleles of each contributor in a mixture can become hard to achieve even in simpler cases such as a mixed profile from two contributors. As a consequence of these new challenges, forensic laboratories, and the database managers may use various criteria to limit the deposition of mixed or partial profiles into the NDDB. For instance, the NDDB will only accept mixtures with data for L STR loci, where $9 \leq L \leq 13$ with a maximum number of loci exhibiting more than two alleles equal to $L-7$, and with no more than five alleles per locus. Although STRs exhibit very high level of polymorphism enabling high discriminatory power, they are subject, like any amplification-based markers, to the presence of polymorphisms within the primer binding site which results in lack of amplification or so-called drop-out alleles (or null alleles). The impact of such result has been well documented (Haned et al., 2011) and probabilistic

methods can be used to account for drop-in and drop-out alleles (Gill et al., 2012).

With these managerial constraints, the development of statistical methodologies allowing more formal quantitative comparisons of casework profiles to DNA databanks is required. In the meantime, the LSJML has developed an innovative investigative strategy to increase the use of partial profiles from LCN or complex mixtures in their search for matches in the databanks, relying on two complementary practices.

Elaborated interpretation and databanking guidelines at LSJML allow the specific extraction of the relevant genetic information contained in single-source or mixed profiles for databank searches for intelligence purposes (Noël et al., 2009). For instance, the flagging of alleles as “obligate” or “non-obligate” in queries sent to the NDDB allows filtering out considerably the potential matches, limiting them to a subset of possible matches that is consistent (see section Challenges of the LSJML model and research prospects) with all the information available for the casework. For example, this procedure is used to separate alleles that are likely to come from the putative aggressor in intimate swabs from the victim of a sexual assault—i.e., alleles that must be included in any candidate match returned by the NDDB—from alleles of less certain origin (e.g., alleles of the victim potentially shared with the aggressor) that need not be present in the candidate profile. More generally, this approach is valid for any mixture related to any type of infraction where some of the alleles are more likely than others to come from the offender(s). Another option is to eliminate alleles from a person whose DNA profile is known from other traces obtained for the same casework (e.g., victim, witness or single-source unknown), or those that would imply either highly unbalanced peak heights of a contributor to a mixture or dropouts when it is not a reasonable possibility based on statistical data. It is up to the reporting scientist to check the relevance of the hypothesis with his/her scientific investigation of the case.

The second aspect of the LSJML strategy is the maintenance of its own local database (also hosted in the CODIS system) where complex mixtures that do not meet the NDDB criteria can be deposited, namely in the “Forensic High Mixture” index, for comparison with other local casework profiles (Figures 1, 2). In addition, the local database allows searching for matches using more loci, i.e., up to 15 at the LSJML operational setup instead

of the 13 CODIS loci in the NDDB. Finally, mixed strategies are authorized whereby a full mixture can be deposited into the local database while a subset of its alleles (a “submixture”) is sent to the NDDB. Thus, matches can potentially occur at the local level between the whole mixture kept as a “backup” and pure or mixed profiles from other caseworks. This can be especially useful when deconvolution is difficult so that there is much uncertainty around which alleles should be sent at the NDDB.

CHALLENGES OF THE LSJML MODEL AND RESEARCH PROSPECTS

While the LSJML model provides great flexibility in maximizing the number of matches, it also raises legitimate questions about potential biases that may arise from its databanking strategy (Lynch, 2003; Dror et al., 2006; Dror and Hampikian, 2011).

Aware of it, the LSJML has adopted different strategies to assess their importance and limit them. These range from operational rules to current and prospective research projects. First, the LSJML does not declare a match as valid as soon it occurs (except when both the target and the candidate are single-sourced and complete). Thus, once a match between a target profile and a candidate profile in the NDDB has occurred, the LSJML scientist must assess its validity. The procedure involves an evaluation of the candidate profile using the original electropherogram from which the target profile was extracted, statistical data on peak height balance and drop-outs, as well as other profiles from the casework. This is the step where consistence with all the information available for the casework is evaluated. In addition, the validity of the match must also be confirmed by one of the two local scientists managing the databank.

Second, because the above procedure may limit but not completely eliminate fortuitous (wrong) matches, the opinion on evidential weight (Providers, 2009) is based on standard statistical approaches such as the probability of exclusion or likelihood ratios performed on the whole mixture, and not on extracted elements, except when a major profile can clearly be extracted using strict deconvolution rules.

Third, the LSJML has been proactive in challenging the validity of its own strategy with respect to biases or invalid match generation by undertaking a number of quantitative statistical evaluations. It is worthy to note that the databanking and match review strategies for targets arising from mixtures of various levels of

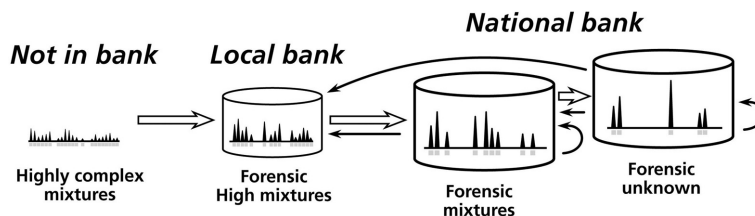


FIGURE 2 | The processes of mixture databanking and comparison at LSJML. The complexity of profiles decreases from left to right, i.e., from highly complex mixtures that cannot be deposited as is in databank to single-source profiles stored under the “forensic unknown” index. “Forensic high mixture” and “Forensic mixture” are two intermediate indexes, respectively stored at the local (LSJML) and national (NDDB)

levels. These three indexes composed the Crime Scene Index (CSI; see Figure 1). Open arrows show how a mixture can switch category when alleles are removed from it (e.g., alleles of low intensity or from a known contributor; see section The LSJML databanking strategy). Solid arrows indicate how profiles from the different indexes are compared in search for matches.

complexity generate valid candidate matches in comparable proportions to single-source targets (for which match validity is automatic), with similar levels of effort (i.e., working time required to evaluate the matches) see (Noël et al., 2009) and (Lavergne et al., 2008) for details. For instance, less than 12% of candidate matches produced with the Forensic Unknown and Forensic Mixture indexes (**Figure 2**) are rejected with a “no match” disposition after review. One critical aspect is that mixtures must be of good quality, namely show good peak intensities. Moreover, the LSJML has begun to perform experimental tests by searching two-person mixtures with up to 13 mixed loci against the Florida data bank constituted of nearly 500,000 convicted offender profiles. Because of the geographical (~2000 km) and country barriers between Québec and Florida, it is expected that almost any eventual match would be fortuitous. Corroborating above conclusions, these mixtures did not return more candidate matches than less complex ones. Moreover, all candidate matches were rejected independently (i.e., not in concert) by four reporting scientists. Finally, the lab, in collaboration with others, is presently evaluating an alternative to the current selection procedure for uploading mixtures to the NDDB. The new approach would be based on the number of expected matches accounting for the COI size and is implemented in the CODIS Match Estimator® module.

At this time, open discussion between LSJML and academic partners to better assess the potential hazards of inducing these databanking policies with respect to confirmations bias are currently underway. Nevertheless, an understanding of this strategy with respect to a possible future goal toward forensic intelligence should be kept in mind (Ribaux et al., 2006; Pham-Hoi et al., 2013). On the other hand, limiting decisions into whether identification was correct or not by only using pure profiles may provide a sense of security. However, this also leads to the restricted use of the information available, with potentially pertinent information discarded when solving everyday crimes. It is currently unclear what the consequences of refusing to tackle these issues will have on victims, and consequently on justice, who also has a validating role to play in this area.

Nevertheless, a fine-tuned approach, specific to the various types of casework (sexual assault, homicide, burglary, high-volume crimes, etc.) definitely needs to be addressed to better, and more rigorously, assess the consequences these changes will have on the whole process of identification.

BEYOND THE PRESENT DNA PRACTICE

Notwithstanding these innovative practices and the relevant interpretation process to be developed being a sign of academic-practitioner joint effort, the development of STR mixture analysis and databank searching will eventually reach its limit impeding further improvements owing to the inherent limitations of using small sets of markers (typically < 20 for STR) typed by technologies that do not permit to separate DNA from different cells found in the same trace (with the exception of differential extraction of semen DNA). Ultimately, substantial increase in the power of mixture analysis will come from newer technologies such as single nucleotide polymorphisms (SNPs) (Daniel and Walsh, 2006; Kidd et al., 2006; Sanchez et al., 2006; Fang et al., 2009; Pakstis et al.,

2010; Voskoboinik and Darvasi, 2011) or single-cell sequencing (Hanson and Ballantyne, 2005). Repositories like the NDDB will need to adapt to these forthcoming innovations in a way that permit forensic labs to benefit from the full power of these new tools for match searching, but without compromising on the usefulness of the STR information accumulated since their creation.

Other advances in the biological sciences, not strictly depending on the NDDB itself, could benefit from the advice/input/review of the Advisory committee to pave the way for a new forensic dimension (Daniel and Walsh, 2006; Kidd et al.,

Box 1 | Beyond present DNA practices

1) Going further with genomics

While it was initially believed PCR would be capable of solving the challenges encountered from analysing LCN samples, difficulties associated with interpretation lessens its pragmatic use in forensic casework (Gill et al., 2000; Kloosterman and Kersbergen, 2006; McCartney, 2008; Budowle et al., 2009) (see also LCN DNA Review at <http://www.mccannfiles.com/id190.html>). In addition, the use of a limited set of markers, as is the case today, restricts the potential discriminative power that could be accessed if using full genome sequences.

Single-cell genome sequencing is a rapidly improving technology with one of many applications including the detection of somatic intra-individual variation in cancer patients (Navin et al., 2011). Initially applied to small prokaryotic genomes (Stepanaukas and Sieracki, 2007), recent advances in next-generation sequencing have enabled the coverage of 93% of the much larger human genome from a single human cell (Zong et al., 2012). This not only allows for the identification of SNPs and LCN variation (Zong et al., 2012) but also genomic structural variation and somatic mutations that give rise to intra-individual genetic variation (O'Huallachain et al., 2012). With as many as 2500 genomic structural variations and three million SNPs (Abecasis et al., 2010) occurring between two unrelated individuals, the discriminative power of this technique could even allow for identical twins to be differentiated.

Full genome sequencing would permit the use of a much wider range of genomic polymorphism to convict or exonerate persons of interest. Although currently cost prohibitive, and against the current ideology that the use of anonymous loci is preferable, the dwindling cost associated with genome sequencing may enable their use in a foreseeable future.

2) Adding transcriptomics and proteomics to the forensic toolbox ?

With research carried out in the fields of transcriptomics and proteomics, opportunities are emerging to develop and add to the already existing genomic platforms.

Evidence of physical abuse is often left of the skin of victims, with bruising found to be the most common form of injury (Dye et al., 2008; Pierce et al., 2010; Jackson et al., 2012). The ability to accurately and reliably determine the age of a bruise, in living individuals, is currently lacking. If possible, this could provide vital evidence to legal cases of suspected physical abuse. In cases where multiple bruises are present on the body of a victim, providing evidence that the injuries were inflicted on separate occasions could have important medico-legal significance.

Building a human proteome map of protein markers present in unbruised skin, as well as bruised skin, and analysing changes in protein expression levels as a bruise evolves, could help to achieve these goals (Lecomte et al., 2013).

2006). Research into fields such as ancestry informative markers (AIMs) (Lao et al., 2008; Kersbergen et al., 2009; Kosoy et al., 2009; Liu et al., 2009), proteomics (Kool et al., 2007; Lecomte et al., 2013), genome/marker based phenotyping (Sulem et al., 2008; Liu et al., 2009; Zubakov et al., 2010; Walsh et al., 2011), framing the input of DNA to forensic intelligence (Jobling and Gill, 2004; Ribaux et al., 2006; Bond, 2007; Roman et al., 2009; Wilson et al., 2011), and the incoming lab-on-a-chip involvement of crime scene (Batt et al., 2009; Bell, 2011). All these fields belong to a still-debated investigative process (Kaye, 2007) opposed to the claim for a strict separation of laboratories from the law enforcement system (Nrc, 2009). **Box 1** presents two examples of techniques that could eventually be used in forensic sciences on a case-by-case basis. One of them, forensic proteomics, does not directly assist to the evolution of the NDDB. However, the power of these new tools to address personal characteristics of human beings, could lead to an ethical position being taken by the Advisory Committee, which could impact the future developments of the data bank.

CONCLUSION

As exciting projects make their way in the field of molecular biology, real challenges also lie in the realm of forensic science, giving new impedimenta to forensic DNA and, raising obvious ethical, social, and economic questions. Nevertheless, the inescapable drive toward DNA intelligence and laboratory miniaturization, and the projection on the crime scene, could underline the need for a better scientific support of the crime scene officers present at the start of the forensic process.

As commissioner Paulson of the NDDB wrote in the last annual report, “the NDDB operates within a diverse environment that must consider scientific advancements, privacy rights, and changing legislation.” In regards to the building up of the NDDB and the wisdom of its Advisory committee, an optimistic future for the scientific support of the Canadian law and justice systems is anticipated.

AUTHOR CONTRIBUTION

Emmanuel Milot is post-PhD researcher in biology and consultant for the LSJML and the RCMP laboratories. His research interest addresses two domains: the use of genetics to study human and animal population dynamics and the causes of phenotypic variations between individuals.

Marie M. J. Lecomte has recently submitted her PhD employing proteomic techniques to determine the age of bruises in living individuals.

Hugo Germain is professor of biochemistry, head of chair on vegetal immunity, in charge of human DNA identification course at the forensic curriculum at the UQTR.

Frank Crispino is criminalist with a PhD degree from the University of Lausanne (Switzerland) and a former criminal and counter-terrorist investigation commander in the French Gendarmerie. He is now serving as professor in forensic science at the UQTR.

Although this article is a common elaborated paper, the order of appearance reflects their relative inputs in the paper, the corresponding author being, moreover, in charge of coordination.

ACKNOWLEDGMENTS

The authors would like to thank M. Tony Tassarolo, director of the Center of Forensic Sciences, Ontario Province, for his kind highly appreciated relevant comments on a draft version of this paper and Ron Fournery from the NDDB for insightful discussions. We are also indebted to Josée Noël and Léo Lavergne from the Laboratoire de sciences judiciaires et de médecine légale for contributing important information to this paper and commenting the draft. We also thank Mireille Courteau who kindly drew the final version of the figures. Finally, we express our gratitude to the two unknown reviewers who helped us upgrade the original draft.

REFERENCES

- Abecasis, G. R., Altshuler, D., Auton, A., Brooks, L. D., Durbin, R. M., Gibbs, R. A., et al. (2010). A map of human genome variation from population-scale sequencing. *Nature* 467, 1061–1073. doi: 10.1038/nature09534
- Batt, C. A., Stelick, S. J., Kennedy, M. J., Lui, C. S., and Lowe, A. J. (2009). *A Hand-Held DNA-Based Forensic Tool*, ed U. D. O. Justice, Rockville, MD: National Institute of Justice.
- Bell, S. (2011). “Forensic microfluidics outside the DNA box,” in *15th International Conference on Miniaturized Systems for Chemistry and Life Sciences*. (Seattle, WA).
- Bieber, F. R., Brenner, C. H., and Lazer, D. (2006). Human genetics—finding criminals through DNA of their relatives. *Science* 312, 1315–1316. doi: 10.1126/science.1122655
- Bill, M., Gill, P., Curran, J., Clayton, T., Pinchin, R., Healy, M., et al. (2005). PENDULUM—a guideline-based approach to the interpretation of STR mixtures. *Forensic Sci. Int.* 148, 181–189. doi: 10.1016/j.forsciint.2004.06.037
- Bond, J. W. (2007). Value of DNA evidence in detecting crime. *J. Forensic Sci.* 52, 128–136. doi: 10.1111/j.1556-4029.2006.00323.x
- Budowle, B., Eisenberg, A., and van Daal, A. (2009). Validity of low copy number typing and applications to forensic science. *Croat. Med. J.* 50, 207–217. doi: 10.3325/cmj.2009.50.207
- Curran, J. M. (2008). A MCMC method for resolving two person mixtures. *Sci. Justice* 48, 168–177. doi: 10.1016/j.scijus.2007.09.014
- Curran, T. (1997). “Forensic DNA analysis: technology and applications,” in *Background Paper BP-443E*, ed Bibliothèque du parlement (Ottawa, ON: Parliamentary Research Branch, Library of Parliament), 43. available online at <http://www.parl.gc.ca/Content/LOP/researchpublications/bp443-e.htm>
- Daniel, R., and Walsh, S. J. (2006). The continuing evolution of forensic DNA profiling—from STRs to SNPs. *Aust. J. Forensic Sci.* 38, 59–74. doi: 10.1080/00450610609410633
- Dror, I. E., Charlton, D., and Peron, A. (2006). Contextual information renders experts vulnerable to making erroneous identifications. *Forensic Sci. Int.* 174–178. doi: 10.1016/j.forsciint.2005.10.017
- Dror, I. E., and Hampikian, G. (2011). Subjectivity and bias in forensic DNA mixture interpretation. *Sci. Justice* 51, 204–208. doi: 10.1016/j.scijus.2011.08.004
- Dye, D. W., Peretti, F. J., and Kokes, C. P. (2008). Histologic evidence of repetitive blunt force abdominal trauma in four pediatric fatalities. *J. Forensic Sci.* 53, 1430–1433. doi: 10.1111/j.1556-4029.2008.00883.x
- Fang, R., Pakstis, A. J., Hyland, E. C., Wang, D., Shewale, J., Kidd, J. R., et al. (2009). Multiplexed SNP detection panels for human identification. *Forensic Sci. Int. Genet.* 2, 538–539. doi: 10.1016/j.fsigss.2009.08.161
- Foreman, L. A., Champod, C., Evett, I. W., Lambert, J. A., and Pope, S. (2003). Interpreting DNA evidence: a review. *Int. Stat. Rev.* 71, 473–495. doi: 10.1111/j.1751-5823.2003.tb00207.x
- Gershaw, C. J., Schweighardt, A. J., Rourke, L. C., and Wallace, M. M. (2011). Forensic utilization of familial searches in DNA databases. *Forensic Sci. Int. Genet.* 5, 16–20. doi: 10.1016/j.fsigen.2010.07.005
- Gill, P., Whitaker, J., Flaxman, C., Brown, N., and Buckleton, J. (2000). An investigation of the rigor of interpretation rules for STRs derived from less than

- 100 pg of DNA. *Forensic Sci. Int.* 112, 17–40. doi: 10.1016/S0379-0738(00)00158-4
- Gill, P., Gusmao, L., Haned, H., Mayr, W. R., Morling, N., Parson, W., et al. (2012). DNA commission of the international society of forensic genetics: recommendations on the evaluation of STR typing results that may include drop-out and/or drop-in using probabilistic methods. *Forensic Sci. Int. Genet.* 5, 679–688. doi: 10.1016/j.fsigen.2012.06.002
- Haned, H., Egeland, T., Pontier, D., Pene, L., and Gill, P. (2011). Estimating drop-out probabilities in forensic DNA samples: a simulation approach to evaluate different models. *Forensic Sci. Int. Genet.* 5, 525–531. doi: 10.1016/j.fsigen.2010.12.002
- Hanson, E. K., and Ballantyne, J. (2005). Whole genome amplification strategy for forensic genetic analysis using single or few cell equivalents of genomic DNA. *Anal. Biochem.* 346, 246–257. doi: 10.1016/j.ab.2005.08.017
- Jackson, J., Carpenter, S., and Anderst, J. (2012). Challenges in the evaluation for possible abuse: presentations of congenital bleeding disorders in childhood. *Child Abuse Negl.* 36, 127–134. doi: 10.1016/j.chiabu.2011.09.009
- Jobling, M. A., and Gill, P. (2004). Encoded evidence: DNA in forensic analysis. *Nat. Rev. Genet.* 5, 739–751. doi: 10.1038/nrg1455
- Kaye, D. H. (2007). The science of DNA identification: from the laboratory to the courtroom (and Beyond). *Minn. J. L. Sci. Tech.* 8, 409–427.
- Kayser, M., and de Knijff, P. (2011). Improving human forensics through advances in genetics, genomics and molecular biology. *Nat. Rev. Genet.* 12, 179–192. doi: 10.1038/nrg2952
- Kersbergen, P., van Duijn, K., Kloosterman, A. D., den Dunnen, J. T., Kayser, M., and de Knijff, P. (2009). Developing a set of ancestry-sensitive DNA markers reflecting continental origins of humans. *BMC Genet.* 10:69. doi: 10.1186/1471-2156-10-69
- Kidd, K. K., Pakstis, A. J., Speed, W. C., Grigorenko, E. L., Kajuna, S. L., Karoma, N. J., et al. (2006). Developing a SNP panel for forensic identification of individuals. *Forensic Sci. Int.* 164, 20–32. doi: 10.1016/j.forsciint.2005.11.017
- Kloosterman, A. D., and Kersbergen, P. (2003). Efficacy and limits of genotyping low copy number DNA samples by multiplex PCR of STR loci. *Int. Congr. Ser.* 1239, 795–798. doi: 10.1016/S0531-5131(02)00514-9
- Kool, J., Reubsaet, L., Wesseldijk, F., Maravilha, R. T., Pinkse, M. W., D'Santo, C. S., et al. (2007). Suction blister fluid as potential body fluid for biomarker proteins. *Proteomics* 7, 3638–3650. doi: 10.1002/pmic.200600938
- Kosoy, R., Nassir, R., Tian, C., White, P. A., Butler, L. M., Silva, G., et al. (2009). Ancestry informative marker sets for determining continental origin and admixture proportions in common populations in america. *Hum. Mutat.* 30, 69–78. doi: 10.1002/humu.20822
- Lao, O., Lu, T. T., Nothnagel, M., Junge, O., Freitag-Wolf, S., Caliebe, A., et al. (2008). Correlation between genetic and geographic structure in Europe. *Curr. Biol.* 18, 1241–1248. doi: 10.1016/j.cub.2008.07.049
- Lavergne, L., Mailly, F., Noël, J., and Jolicoeur, C. (2008). “Searching a DNA data bank using complex mixtures; a retrospective study,” in *19th International Symposium on Human Identification*. (Hollywood, CA).
- Lecomte, M. M. J., Atkinson, K. R., Kay, D. P., Simons, J. L., and Ingram, J. R. (2013). A modified method using the SonoPrep® ultrasonic skin permeation system for sampling human interstitial fluid is compatible with proteomic techniques. *Skin Res. Tech.* 19, 27–34. doi: 10.1111/j.1600-0846.2012.00641.x
- Liu, F., van Duijn, K., Vingerling, J. R., Hofman, A., Uitterlinden, A. G., Janssens, A. C. J. W., et al. (2009). Eye color and the prediction of complex phenotypes from genotypes. *Curr. Biol.* 19, R192–R193. doi: 10.1016/j.cub.2009.01.027
- Lowe, A., Murray, C., Whitaker, J., Tully, G., and Gill, P. (2002). The propensity of individuals to deposit DNA and secondary transfer of low level DNA from individuals to inert surfaces. *Forensic Sci. Int.* 129, 25–34. doi: 10.1016/S0379-0738(02)00207-4
- Lynch, M. (2003). God's signature: DNA profiling, the new gold standard in forensic evidence. *Endeavor* 27, 93–97. doi: 10.1016/S0160-9327(03)00068-1
- McCartney, C. (2008). LCN DNA: proof beyond reasonable doubt? *Nat. Rev. Genet.* 9:325. doi: 10.1038/nrg2362
- Meyers, S. P., Timken, M. D., Piucci, M. L., Sims, G. A., Greenwald, M. A., Weigand, J. J., et al. (2011). Searching for first-degree familial relationships in California's offender DNA database: validation of a likelihood ratio-based approach. *Forensic Sci. Int. Genet.* 5, 493–500. doi: 10.1016/j.fsigen.2010.10.010
- Miller, G. (2010). Familial DNA testing scores a win in serial killer case. *Science* 329, 262. doi: 10.1126/science.329.5989.262
- Murphy, E. (2010). Relative doubt: familial searches of DNA databases. *Mich. L. Rev.* 109, 291–348.
- Navin, N., Kendall, J., Troge, J., Andrews, P., Rodgers, L., McIndoo, J., et al. (2011). Tumour evolution inferred by single-cell sequencing. *Nature* 472, 90–94. doi: 10.1038/nature09807
- Noël, J., Lavergne, L., Mailly, F., Roberge, D., and Jolicoeur, C. (2009). Searching a DNA databank with complex mixtures from two individuals. *Forensic Sci. Int. Genet.* 2, 464–465. doi: 10.1016/j.fsigs.2009.08.119
- Nrc. (2009). *Strengthening Forensic Science in the United States: A Path Forward*. Washington, DC: The National Academies Press.
- O'Huallachain, M., Karczewski, K. J., Weissman, S. M., Urban, A. E., and Snyder, M. P. (2012). Extensive genetic variation in somatic human tissues. *Proc. Natl. Acad. Sci. U.S.A.* 109, 18018–18023. doi: 10.1073/pnas.1213736109
- Pakstis, A. J., Speed, W. C., Fang, R., Hyland, F. C., Furtado, M. R., Kidd, J. R., et al. (2010). SNPs for a universal individual identification panel. *Hum. Genet.* 127, 315–324. doi: 10.1007/s00439-009-0771-1
- Pham-Hoi, E., Crispino, F., and Hampikian, G. (2013). The first successful use of a low stringency familial match in a french criminal investigation. *J. Forensic Sci.*
- Phipps, M., and Petricevic, S. (2007). The tendency of individuals to transfer DNA to handled items. *Forensic Sci. Int.* 168, 162–168. doi: 10.1016/j.forsciint.2006.07.010
- Pierce, M. C., Kaczor, K., Aldridge, S., O'Flynn, J., and Lorenz, D. J. (2010). Bruising characteristics discriminating physical child abuse from accidental trauma. *Pediatrics* 125, 67–74. doi: 10.1542/peds.2008-3632
- Police, R. C. M. (2012). “The national DNA data bank of Canada,” in *Annual Report*, ed Forensic Science and Identification Services, Royal Canadian Mounted Police (Ottawa, ON), 1–30.
- Providers, A. O. F. S. (2009). Standards for the formulation of evaluative forensic science expert opinion. *Sci. Justice* 49, 161–164. doi: 10.1016/j.scijus.2009.07.004
- Reid, T. M., Baird, M. L., Reid, J. P., Lee, S. C., and Lee, R. F. (2008). Use of sibling pairs to determine the familial searching efficiency of forensic databases. *Forensic Sci. Int. Genet.* 2, 340–342. doi: 10.1016/j.fsigen.2008.04.008
- Ribaux, O., Walsh, S. J., and Margot, P. (2006). The contribution of forensic science to crime analysis and investigation: forensic intelligence. *Forensic Sci. Int.* 156, 171–181. doi: 10.1016/j.forsciint.2004.12.028
- Roman, J. K., Reid, S., Reid, J., Chalfin, A., Adams, W., and Knight, C. (2009). *The DNA Field Experiment: Cost-Effectiveness Analysis of the Use of DNA in the Investigation of High-Volume Crimes*, (Washington, DC: Urban Institute).
- Sanchez, J. J., Phillips, C., Borsting, C., Balogh, K., Bogus, M., Fondevila, M., et al. (2006). A multiplex assay with 52 single nucleotide polymorphisms for human identification. *Electrophoresis* 27, 1713–1724. doi: 10.1002/elps.200500671
- Stepanaukas, R., and Sieracki, M. E. (2007). Matching phylogeny and metabolism in the uncultured marine bacteria, one cell at a time. *Proc. Natl. Acad. Sci. U.S.A.* 104, 9052–9057. doi: 10.1073/pnas.0700496104
- Sulem, P., Gudbjartsson, D. F., Stacey, S. N., Helgason, A., Rafnar, T., Jakobsdottir, M., et al. (2008). Two newly identified genetic determinants of pigmentation in Europeans. *Nat. Genet.* 40, 835–837. doi: 10.1038/ng.160
- Voskoboinik, L., and Darvasi, A. (2011). Forensic identification of an individual in complex DNA mixtures. *Forensic Sci. Int. Genet.* 5, 428–435. doi: 10.1016/j.fsigen.2010.09.002
- Walsh, S., Lindenbergh, A., Zuniga, S. B., Sijen, T., de Knijff, P., Kayser, M., et al. (2011). Developmental validation of the IrisPlex system: determination of blue and brown iris color for forensic intelligence. *Forensic Sci. Int. Genet.* 5, 464–471. doi: 10.1016/j.fsigen.2010.09.008
- Walsh, S. J. (2009). *Evaluating the Role and Impact of Forensic DNA Profiling on Key Areas of the Criminal Justice System*. Sydney: University of technology of Sydney.

- Wilson, D. B., Weisburd, D., and McClure, D. (2011). Use of DNA testing in police investigative work for increasing offender identification, arrest, conviction and case clearance. *Campbell Syst. Rev.* 7, 1–56. doi: 10.4073/csr.2011.7
- Zong, C., Lu, S., Chapman, A. R., and Xie, X. S. (2012). Genome-wide detection of single-nucleotide and copy-number variations of a single human cell. *Science* 338, 1622–1626. doi: 10.1126/science.1229164
- Zubakov, D., Liu, F., van Zelm, M. C., Vermeulen, J., Oostra, B. A., van Duijn, C. M., et al. (2010). Estimating human age from T-cell DNA rearrangements. *Curr. Biol.* 20, R970–R971. doi: 10.1016/j.cub.2010.10.022

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 16 August 2013; paper pending published: 22 September 2013; accepted: 31 October 2013; published online: 20 November 2013.

Citation: Milot E, Lecomte MMJ, Germain H and Crispino F (2013) The National DNA Data Bank of Canada: a Quebecer perspective. *Front. Genet.* 4:249. doi: 10.3389/fgene.2013.00249

This article was submitted to Statistical Genetics and Methodology, a section of the journal *Frontiers in Genetics*.

Copyright © 2013 Milot, Lecomte, Germain and Crispino. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.